

Overview: The U.S. Office of Naval Research Training Technology R&D

Dr. Susan F. Chipman

Office of Naval Research, Code 342
800 N. Quincy Street
Arlington, Virginia 22217-5660
United States of America

Summary

The U.S. Office of Naval Research (ONR) has a long, continuous, and distinguished tradition of support for research and development of advanced technologies for application to military training. This paper provides an overview of the history and structure of those research programs with emphasis on areas of major current activity: 1) basic research on tutorial dialog aiming for true natural language interaction capability for artificially intelligent training systems; 2) effective instructional strategies for artificially intelligent coaching in dynamically evolving situations such as instrument flying, shipboard command information center tasks, and shipboard damage control and 3) the integration of artificially intelligent coaches into virtual reality simulators. The content of this paper is very closely related to material available on an ONR web site designed for do-it-yourself briefings : (http://www.onr.navy.mil/sci_tech/personnel/cnb_sci/342/default.htm).

*Paper presented at the RTO HFM Symposium on "Advanced Technologies for Military Training",
held in Genoa, Italy, 13 – 15 October 2003, and published in RTO-MP-HFM-101.*

Report Documentation Page			<i>Form Approved OMB No. 0704-0188</i>					
<p>Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p>								
1. REPORT DATE 00 APR 2004	2. REPORT TYPE N/A	3. DATES COVERED -						
4. TITLE AND SUBTITLE Overview: The U.S. Office of Naval Research Training Technology R&D			5a. CONTRACT NUMBER					
			5b. GRANT NUMBER					
			5c. PROGRAM ELEMENT NUMBER					
6. AUTHOR(S)			5d. PROJECT NUMBER					
			5e. TASK NUMBER					
			5f. WORK UNIT NUMBER					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Office of Naval Research, Code 342 800 N. Quincy Street Arlington, Virginia 22217-5660 United States of America			8. PERFORMING ORGANIZATION REPORT NUMBER					
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)					
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)					
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release, distribution unlimited								
13. SUPPLEMENTARY NOTES See also ADM001667, NATO RTO-MP-HFM-101 Advanced Technologies for Military Training (Technologies avancées pour l'entraînement militaire)., The original document contains color images.								
14. ABSTRACT								
15. SUBJECT TERMS								
16. SECURITY CLASSIFICATION OF: <table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 33%; padding: 2px;">a. REPORT unclassified</td> <td style="width: 33%; padding: 2px;">b. ABSTRACT unclassified</td> <td style="width: 33%; padding: 2px;">c. THIS PAGE unclassified</td> </tr> </table>			a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified	17. LIMITATION OF ABSTRACT UU	18. NUMBER OF PAGES 20	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified						

This NATO Symposium on Advanced Technologies for Military provided a very appropriate opportunity to review the important contributions to advanced training technologies that the U.S. Office of Naval Research (ONR) has made over the years, since it was founded in 1946 to maintain the relationships with university researchers that had proved very useful during World War II. Few people recognize how important research and development support from ONR, as well as other U.S. Department of Defense research agencies, has been to the development of education and training technologies in the United States. Research in psychology, including research directed at training applications, was included in the research programs from the outset, so ONR has had a long, continuous, and distinguished tradition of support for advanced training technologies. Consequently the subject of this paper is a very large one that could easily fill an entire book. Here it is possible to provide only a rather brief overview. As of this writing, however, more information can be found on a public ONR web site that it designed for do-it-yourself briefings adapted to the interests of the user. Its current url is:

(http://www.onr.navy.mil/sci_tech/personnel/cnb_sci/342/default.htm).

ONR's Strong History of Contributions to Training Technology

In 1993, I was the invited speaker at the last meeting of the Military Special Interest Group of the American Educational Research Association, asked to speak on the contributions of DoD to education and training technology. At that time, I reviewed the complete collection of annual project books, describing the research projects that ONR had supported. (Unfortunately these books are no longer being produced.) These were quite revealing. In the late 1950's and the 1960's, ONR supported all the well-known pioneers of computer-assisted or computer-based training as we know it today. These efforts focused on the goal of individualized instruction, including mastery testing, branching instruction based on testing outcomes, and selective remedial instruction. Even today, this sort of individualization is not as frequently implemented in computer-based instruction as we would hope. In the 1970's investigations of what we could call interactive video began. The first investigations predated the invention of the interactive videodisk, using picture presentation devices designed for psychological experiments in animal learning. An early study demonstrated that pictures simulating interaction with a piece of equipment could be as effective as interaction with the real equipment in maintenance training applications. This was the beginning of a line of work described by Towne (1987). Support of research on issues of effective multimedia design, initially in the videodisk environment, began about 1980, notably the research of Patricia Baggett. Her research is summarized in Wetzel, Radtke, and Stern (1994), *The Instructional Effectiveness of Video Media*, a broad review that is itself a product of Navy training research investments.

A major emphasis of ONR's research programs over the years has been artificially intelligent tutoring, also known as intelligent computer-assisted instruction (ICAI). The first contract in this area was awarded to the late Jaime Carbonell of Bolt, Beranek and Newman, Inc., in 1969. At that time, a \$2M computer was required to do the research. In the 1980's intelligent tutoring became an important program emphasis with a substantial increase in dedicated funding. During part of this period, both the U.S. Army Research Institute (ARI) and the Defense Advanced Research Projects Agency (DARPA) participated in jointly funded projects with ONR. Under the Defense University Research Instrumentation Program, ONR also provided many members of the university research community AI workstations to be used in intelligent tutoring research. By that time, the cost of the necessary computers had come down to \$20-\$60K each. The important role of ONR and other DoD agencies in the development of artificially intelligent tutoring technologies was very much related to the willingness to invest in long term research using computers that were, at the time, too expensive for practical instructional use. In contrast, a joint program of the U.S. National Institute of Education and the National Science Foundation, Science Education Directorate, supported R&D exploring the use of computers in mathematics education at about the same time. That program required that the researchers use computers that could be expected to be in classrooms within five

years, and most of the projects used Apple II computers that actually were in classrooms at the time. Such computers were not adequate to implement artificially intelligent tutoring approaches.

By the time the special funding emphasis on artificially intelligent tutoring concluded, one could fairly say that the American, and international, research community basically knew how to build highly effective tutoring systems. Among the papers and books summarizing the accomplishments of this era are Anderson, Boyle, & Reiser (1985), Clancey (1986), Lesgold, Chipman, Brown, & Soloway, (1990), Wenger, (1987).

There were also two notable projects demonstrating that the technology could apply to practical military maintenance training problems. One of these was the Navy's Intelligent Maintenance Training System (IMTS) (Towne, 1987) that also led to research on authoring tools to build similar systems at lower costs. The other was the Air Force's SHERLOCK (Gott, 1989), which involved one of the best evaluations of the effectiveness of intelligent tutoring that has yet been done. Although SHERLOCK was an Air Force project, it built upon earlier ONR sponsored work both at the University of Pittsburgh and Bolt, Beranek and Newman. In addition, ONR funded a demonstration application to troubleshooting ship steam power plants (Vasandani, 1991). The National Science Foundation also began to support applications of artificially intelligent tutoring to education, mainly supporting researchers whose earlier work had been supported by ONR (Koedinger, Anderson, Hadley, & Mark (1995), Schofield, (1995).

In the 1990's, ONR, along with many other agencies, lept on the virtual reality band wagon. Substantial investments were made. Some relatively basic research was supported, investigating such issues as simulator sickness in virtual reality. As the decade closed, some practical training applications for the Navy, that will be described below, were beginning to emerge. Emphasis is now shifting from trainers of vehicle operation for both the Navy and Marine Corps to applications in urban warfare for the Marines.

Investments related to artificially intelligent tutoring continue. At the basic research level, current efforts aim at achieving true natural language interaction capability for artificially intelligent tutoring systems, tutorial dialog. Cognitive modeling capability is also being expanded, notably into the psychological representation of space and spatial reasoning abilities. This will expand the range of model-based tutoring systems that can be built. Somewhat more applied research projects combine investigations of instructional strategy issues with building demonstration tutors.

The Vital Importance of DoD Investment to the Advance of Training Technology

Unfortunately, our ultimate military customers for training technology do not always realize that the technology they see industry offering them today probably would not exist if the underlying research had not been supported by the DoD agencies many years before. Many presume that the technology was actually developed by industry or that the Department of Education takes care of that sort of thing. For this reason, the following quotation from Frances Degen Horowitz, as she left her position as the American Psychological Association's Chief Science Advisor, is one of my favorites. Over the years, she had learned that, "The value of the military as an apolitical setting in which to develop state-of-the-art teaching and testing devices and materials is a factor that ought to be taken into account as plans are developed for the downsizing of the military. The Defense Department funds and applies more research in teaching, learning and testing than does any other branch of the federal government." As a developmental psychologist, Horowitz herself never received DoD funding, but she had learned of its importance over the years, especially through her administrative responsibilities. *Apolitical* is the key word in Horowitz's statement. Usually there is strong consensus about what the goals of military training courses are; the same cannot always be said for civilian education. Consequently, the military has provided a more stable environment for pursuing research on instruction. Much of what is viewed as American "educational" research has actually

been supported by the U.S. Department of Defense, notably by ONR. In the United States, education is generally viewed as the responsibility of local communities or of the individual states. Federal level influence on the content of the curriculum tends to be viewed with suspicion. Even the existence of a federal Department of Education is a relatively recent and insecure development. Consequently there has not been a solid political foundation for civilian educational research and development programs in the United States. The U.S. military cannot rely upon the civilian sector as a source for education and training research knowledge.

Artificially Intelligent Tutoring Technology: A Continuing Theme in the ONR Program

The advancement of artificially intelligent tutoring technology has been a major theme of ONR's training research for many years, starting with that first award to Jaime Carbonell of Bolt, Beranek and Newman in 1969. This same underlying technology can be applied in many different delivery options that are sometimes talked about as if they were different technologies: Schoolhouse training using computers, web-based training, shipboard training, embedded training built into military systems. These may look different and may pose different system integration problems, different hardware problems. But the human aspect of the technology, the nature of human learning processes and the instructional interactions that will foster human learning remain much the same.

Why has artificially intelligent tutoring remained a goal for so long? Because individualized tutoring by human tutors seems to be stunningly effective as compared to conventional classroom instruction. Bloom's important paper (Bloom, 1984) argued that it was two standard deviations more effective. However, this could be an underestimate because his evidence included tutoring by inexpert tutors. With a few exceptions, however, individual human tutoring tends to be considered prohibitively expensive, so the pursuit of artificial substitutes becomes very attractive. By now, the enterprise has enjoyed considerable success. A number of artificially intelligent tutors have shown the ability to increase student achievement by about one standard deviation (Anderson, 1993; Koedinger et al., 1995). Although not as many practical tutors as we would like have been built and evaluated, we are beginning to see many topics under the special program for small businesses that ask for the building of artificially intelligent tutors for military use, and a few have been purchased as regular training system procurements (McCarthy, Pacheco, Banta, Wayne, & Coleman, 1994). Artificially intelligent tutoring technology is ready for prime time, and computing technology has matured enough to accommodate it at very reasonable costs. Even production costs for artificially intelligent tutoring systems seem to be comparable to costs for high quality conventional computer-based training. The primary challenge now is to persuade customers for military training systems to ask for this newer technology and to assist them in being smart buyers.

Remaining Challenges for Intelligent Tutoring Research

Although intelligent tutoring research has been quite successful, we still have quite a ways to go to match the effectiveness of human tutors. Here are some of the open research issues that we see in this area:

- True Natural Language Interaction – the special features of tutorial interaction
- Effectiveness of Alternative Instructional Strategies
- Comparative Value of Cognitive Models Differing in Detail/Development Cost
- Overcoming the Cognitive Task Analysis Bottleneck
- Added Value of Sophisticated Tutoring Strategies vs. Simple Reteaching
- Effective use of Multimedia & Virtual Reality

Tutorial Dialog

At the basic research level, the one issue we have chosen to take on at ONR, despite expert advice that it was too difficult, is developing the capacity for artificially intelligent tutoring systems to do true natural language interaction with the student or trainee. Current state-of-the-art tutoring systems are very clever at avoiding the need to do true natural language interaction. It is a salient difference between them and human tutors. Therefore, it seems a promising path for closing the effectiveness gap. Our resources are very limited, so the program goals have focused on understanding student inputs and generating appropriate explanations or other language to present to the student. Neither speech recognition nor speech generation has been included in the program. We have relied on other agencies, notably DARPA, to advance those technologies. While some other agencies have supported work on artificial dialog systems, the focus on the special characteristics of tutorial dialog has been unique to ONR.

Investments began with detailed studies of the linguistic behavior of human tutors (Fox, 1993; Graesser, Person & Magliano, 1995). Next, two computational linguists who also studied human tutorial language but then also attempted to emulate it in computational systems became the key figures in the tutorial dialog program, Martha Evens at the Illinois Institute of Technology and Johanna Moore, who was initially at the University of Pittsburgh but moved to the University of Edinburgh. Over a period of about 10 years, we have made considerable progress. Notably, Martha Evens of the Illinois Institute of Technology produced what I believe to be the first example of a tutor with true natural language interaction capability that reached a level of quality good enough to be used with real students, in this case medical students, who were not being paid to be subjects (Cho, Michael, Rovick, & Evens, 2000). These results were reported immediately after the conference's keynote speaker predicted that such accomplishments were still 10 years in the future. Evens and her associates, as well as Johanna Moore, who is now at Edinburgh, have produced many interesting results about the details of the instructional strategies and linguistic forms that human tutors use (e.g. Hume, Michael, Rovick, & Evens, 1996; Moore, Lemaire & Rosenblum (1996); Core, Moore, & Zinn, (2003)). Evens' project produced a very large number of publications (listed in Evens, 2000), but most of them are in rather obscure conference proceedings. A large number of graduate students (24) also did their dissertation work under this project. Evens and her medical colleagues Michael and Rovick are now writing a book to be published by Erlbaum Associates that will sum up the lessons learned in this research effort. The subject matter domain of the tutor is cardiac physiology; Michael and Rovick are medical school professors, expert tutors who were both the subject of study and collaborators in the effort. Many of the lessons learned about tutoring language and strategies, however, are quite general and already proving very influential. Because Moore had studied human tutoring within the context provided by the SHERLOCK troubleshooting tutor, Moore's early results were used to improve the quality of feedback to trainees in the final delivered version of SHERLOCK, SHERLOCK II. The primary focus of Moore's work is now on developing a clean modularized architecture for tutorial dialog, in which different types of required knowledge are cleanly separated (Zinn, C., Moore, J.D., and Core, M.G., (2002a&b, in press 2003). Another researcher who worked with Moore has now become an independent contributor, Barbara Di Eugenio of the University of Illinois, Chicago (Di Eugenio, 2001).

Tutorial interaction involves gestures as well as speech, and the ONR program has included a small amount of investigation of the instructional use of gestures. Herbert Clark and his students are working with the Stanford MURI project discussed below, investigating gestures. Beth Littleton of Aptima Corporation did some particularly interesting work investigating how instructors and other submariners use gestures while talking about the approach officer's task, which involves very complex spatial reasoning about the localization of another submarine. The use of gestures to express uncertainty was an interesting aspect of the project findings.

Over the last 3 years, the ONR program has been augmented by two large grants with funding from the Office of the Secretary of Defense under the Multi-Disciplinary Research Initiative (MURI) program. One of these projects, directed by Prof. Stanley Peters, a computational linguist at Stanford University, is using an intelligent tutor of shipboard damage control as a demonstration platform (Fry, Ginzton, Peters, Clark, & Pon-Barry (2001). Martha Evens is a consultant to this project. (The damage control tutor, DC-TRAIN, is itself a more applied demonstration project by David Wilkins of the University of Illinois.) Peters has integrated speech recognition technology that was supported in the past by DARPA and Festival text-to-speech technology from the University of Edinburgh with his own natural language technology. DC-TRAIN provides the natural language system with information about what the student did, what errors were made, and what should have been done. The natural language system then provides an after-action review in the form of an interactive dialog with the trainee. A transcript of a demonstration system appears as Appendix A. For this demonstration a domain specific voice was built in Festival by recording quite a number of sentences in the domain: the resulting quality of the speech is so good that only a few small glitches caused by the fact that one word ("do") that was recorded only in a sentence final position reveal that it is indeed computer generated speech. Note also that the system is capable of responding to student initiatives such as the desire not to discuss a particular issue.

The second Tutorial Dialog project, Why2000, tutors qualitative reasoning about physics problems and is a collaboration between Kurt VanLehn at the University of Pittsburgh and Art Graesser at the University of Memphis (Graesser, VanLehn, Rose, Jordan, & Harter, 2001). A number of researchers who worked with Johanna Moore when she was at Pittsburgh are working on this project, as has at least one of Evens' former students. It is exploring combinations of symbolic language processing approaches with more statistical semantic approaches using Latent Semantic Analysis (Landauer, Foltz, & Laham, 1998). It also relates to considerable prior work on physics learning that ONR has supported at the University of Pittsburgh and VanLehn's physics tutor for use at the Naval Academy, described below. At times, you may be able to communicate with the Why2000 tutor over the Internet. It communicates in written language.

Basic Research on Computational Theories of Human Cognitive Architecture

Before we turn to discussion of more applied work in training technologies, the major emphasis of ONR's basic research in Cognitive Science should be mentioned at least briefly. It is integrative computational theories of human cognitive architecture. These theories of cognitive architecture provide theoretical foundations both for artificially intelligent tutoring systems built around detailed cognitive models of the desired knowledge and skill and for cognitive engineering of effective human computer interaction. ONR has been the most important source of support for these integrative theories and supports several of them, notably John Anderson's ACT-R ([Anderson & Lebiere, 1998](#)), which now has a substantial worldwide group of users, Carpenter and Just's CAPS (Just, Carpenter & Varma, 1999; Just & Varma, 2002), and Kieras and Meyer's EPIC (Kieras & Meyer, 1997; Meyer & Kieras, 1997 a & b, 1999). Gentner and Forbus' (Falkenhainer, Forbus, & Gentner, 1989; Forbus, Gentner, & Law, 1995; Gentner, Bowdle, Wolff, & Boronat, 2001) program of research on analogical reasoning and the retrieval of analogical material, with the accompanying computational modeling, has also been supported by ONR. Allen Newell's SOAR (Newell, 1990), which has been more oriented to artificial intelligence than to psychology, has been supported primarily by DARPA, but ONR has supported quite a number of projects in which SOAR was used to model human cognition.

The most important differences among these theories probably are those which arise from the range of phenomena which they have chosen to address first. From very early on, efforts in ACT-R emphasized problem solving and the learning of problem solving of the level of complexity seen in school and college courses. This made it highly relevant to instructional applications. By now, because of the large community of researchers working in the ACT-R framework, a wide range of phenomena have been modeled in ACT-R.

A web site housed at Carnegie Mellon University provides a great deal of information about this work. EPIC set out to account for the major phenomena of so-called human performance research in which people are asked to do more than one simple task simultaneously. Further, Kieras and Meyer adopted an Occam's Razor approach to the development of their architecture: nothing would be assumed in the architecture unless it was actually required in order to account for the empirical data. Although the response delays found in dual-task experiments generally have been used to argue for the existence of central processing limitation, Kieras and Meyer were able to account for the major phenomena in that literature *without* assuming any central processing limitations. They were explained, in precise quantitative detail, by competition for where the eyes are looking and what the hands are doing. These results had major impact on the cognitive modeling research community because they demonstrated the importance of integrating models of perception and motor action with models of cognition. A graduate student quickly explored putting SOAR's "mind" into EPIC's (partial) body (Chong & Laird, 1997), and ACT-R was quickly expanded by incorporating very similar modeling of perceptual and motor systems, becoming ACT-R/PM (Byrne & Anderson, 1997). In contrast, work with CAPS has emphasized situations in which processing limitations are important and serve to explain many individual differences in performance. It is to be expected that EPIC will acquire some central processing limitations when those are required to model some task performance that the research team chooses to address. SOAR was not originally intended to be a valid model of human cognition; it was intended to be a psychologically inspired computational system for work in artificial intelligence. Consequently, it retains a number of unrealistic features as a model of human cognition.

Newer investments emphasize expansion of the range of phenomena that can be integrated into the architectural theory: human representation of space and spatial reasoning, the impact of stress and associated or similar physiological variables on cognitive performance, accounting for the brain activity associated with cognitive activity. These should expand the range of training applications for which high quality cognitive models can be built as the foundation for artificially intelligent training systems. Most of this work is being done in ACT-R because of the growing size of that research community and the relative ease with which new users can learn to work within this theoretical framework due to the training opportunities being provided by Anderson's group with ONR support.

Almost undoubtedly, the reason that military agencies (ONR, DARPA) emerged as the primary supporters of this new theoretical movement was that unified theories of cognition (and perception and motor control) are important to enable practical applications within a reasonable amount of time (say 30 years). Although there are many very solid psychological laws of small scope, it is very difficult to know what they imply about the design of a complex military display or the training of a complex and demanding military task.

Research on Instructional Strategies for Intelligent Tutoring/Demonstration Systems

There are many instructional strategy questions that could, if answered, provide substantial improvements in the effectiveness of artificially intelligent tutoring. Intelligent tutoring systems are very complex systems. Many of the instructional design decisions that must be made have no substantial research basis for guiding them. One must have an intelligent tutoring system in order to investigate the effectiveness of alternative instructional strategies within that context. But building an intelligent tutoring system of realistic scope is too large an investment to be made just for the sake of research investigations. It also remains important to build demonstration intelligent tutoring systems as a strategy to promote the wider application of intelligent tutoring. Consequently, the applied research program (6.2) includes projects which combine the investigation of some generally significant research questions with the building of an intelligent tutoring system that has some value in its own right. Notable examples include a math word problem tutor produced by Sharon Derry of the University of Wisconsin, her students and collaborators (Derry, Wortham,

Webb & Jiang, 1996; Atkinson, 2003) and a coach for physics homework problems developed by Kurt VanLehn at the University of Pittsburgh, in collaboration with several physics professors at the U.S. Naval Academy. Research associated with the math word problem tutor included investigations of how human tutors of adult remedial education characterize individual differences among their students, investigation of the value of using the diagrams from Sandra Marshall's schema theory of math word problem solving (Marshall, 1995; Wortham, 1996)) (also a product of ONR-supported research), and a comparative study of several different approaches for using example problem solutions in instruction (Atkinson, Derry, Renkl & Wortham, 2001; Atkinson, Wortham, Derry, Jiang & Gance, 1998; Wortham, Webb, & Atkinson, 1997). Research with the ANDES physics tutor has also involved comparative evaluation of alternative instructional strategies and efforts to develop an effective strategy for encouraging students to do self-explanation of example problems. The high learning value of self-explanation emerged in earlier ONR-supported work on physics learning (Chi, Bassok, Lewis, Reimann & Glaser, 1989) and has now been replicated several times in a variety of subject matter domains. It is possible that much of the effectiveness of human tutors comes from their ability to support students in developing their own explanations. As is easily done in such systems, the physics tutor records every interaction with the tutor, and the collaborating physics professors have gained many new insights into student problem solving behavior by reviewing the traces of problem solving activity. Relatively late in the process, however, it was discovered that audio recording is a valuable addition for formative evaluation: it revealed student confusion over explanations that had not been detected previously. Successive versions of the physics tutor have been evaluated with classes at the Naval Academy (Albacete & VanLehn, 2000; Shelby, Schulze, Treacy, Wintersgill & VanLehn, 2001). Effect sizes as measured by course exam performance have varied, the largest being .9 standard deviation improvement, but the probable reason for falling short of some past tutors is that the physics professors have been unwilling to implement mastery learning in the tutor because Naval Academy students are so tightly scheduled. We hope that the tutor will be completed soon and made commercially available. There has been considerable interest from the physics education community in general.

The effort to introduce advanced training technology to the Naval Academy also included use of Ken Forbus' CyclePad (Forbus & Whalley, 1994) by Chi Wu, a Naval Academy professor of thermodynamics. Forbus' research on qualitative physics, which underlies the CyclePad software, has been supported for many years by ONR's artificial intelligence program. The development of CyclePad itself was supported by NSF Engineering Education. Supported by ONR, Wu has emerged as the most thorough-going user of CyclePad. He has published many papers on its educational use and is publishing textbooks (Tuttle & Wu, 2001; Wu, 2002, 2003) which reflect revision of course curricula around the use of CyclePad. CyclePad enables students to deal with much more complex and realistic problems. An unclassified version of PC-IMAT (see the paper by Wulfeck) has also been used in instruction at the Naval Academy.

In the last couple of years, a new direction for these tutoring demonstrations involves tutoring in dynamically evolving situations that pose different instructional strategy issues from those that characterize tasks such as programming, or solving math and physics problems, or troubleshooting electronic systems. How does one coach or tutor in these situations without disrupting the performance? John Anderson of Carnegie-Mellon University is working with a one-person simulation of AEGIS anti-air warfare (Sohn, Douglass, Chen, & Anderson, *submitted*), and Stephanie Doane of Mississippi State University is working on turning her successful cognitive model of instrument flying (Doane & Sohn, 2000) into a tutoring system. This tutor involves tracking of eye movements, which are also modeled by the cognitive model and used to assess trainee learning.

An important question about alternative approaches to artificially intelligent tutoring is now being addressed by a contract awarded to James McCarthy of Sonalysts, Inc. under the Small Business Initiative Research (SBIR) Program. The strongest evidence for the effectiveness of intelligent tutoring systems comes from John Anderson's research group. Their tutors are based on extremely fine-grained models of student

knowledge, built as ACT cognitive models. In contrast, other intelligent tutors use coarse grained student models at the level of instructional objectives. Their instructional effectiveness is unknown today. The Sonalysts project will be building two parallel tutors for training Anti-submarine/Anti-surface Tactical Air Controllers (ASTACs), differing only in the grain size of their models. Probably even the fine-grained version will not be quite so fine-grained as the Anderson-style tutors. This issue is of some practical importance because it is obvious that building a fine-grained student model will cost more. In the Sonalysts effort, the cognitive task analysis was conducted in such a way as to provide one data point on the cost difference for the cognitive task analysis; it was not as great as might have been expected.

Authoring Tools for Intelligent Tutoring Systems

Another significant emphasis of the applied research program has been attempts to develop authoring tools that will increase the efficiency of producing such systems and thus reduce the production cost. The Intelligent Maintenance Training System (IMTS) (Towne, 1987) provided a very early example of such tools. Once an author built a simulation of an item of equipment, IMTS could automatically generate an intelligent trouble-shooting tutor by applying a generic space-splitting strategy to the equipment, perhaps modified by available information about the likelihoods of part failure and the costs of investigating various areas of the equipment. IMTS also included tools to facilitate authoring of the simulation itself, as well as various tools for automatically generating instructional routines that were not artificially intelligent. The Air Force later supported a re-implementation of most of the IMTS tools in a PC environment running under Linux, but did not re-implement the capability to generate an intelligent trouble-shooting tutor. That was done under ONR support, with the name DIAG. The DIAG capability was later ported to a Windows environment where it can now be used in conjunction with other Windows-based tools, including graphics tools that now far exceed the capabilities of the IMTS-derived tools. Work with DIAG capabilities now continues in a project funded through the Orlando training lab, "An Integrated Environment for Technical Training and Aiding". This project is exploring widely touted but under-researched concepts for advanced forms of interactive electronic technical manuals that also provide performance aiding and even training functions. They will be working on demonstration applications for the LPD-17, a new ship designed for littoral warfare that will be relying heavily upon computer-based training. Towne has also developed an authoring tool for intelligent scenario-based training. Its demonstration application was to fighting fires in high rise buildings.

Attempts were also made to convert Anderson's laboratory tools for building intelligent tutors to tools that could be used by the wider community. The initial efforts were not successful. However, one of Anderson's former students and collaborators, Ken Koedinger, has continued to work on the "Tutor Development Kit", with support from a variety of sources, including NSF, ONR and DARPA. Increasingly successful summer workshops have been offered for the past few years under NSF support. Of course, we hope that at some future date the current basic research work on tutorial dialog will culminate in authoring tools that will facilitate building true natural language interaction for tutors in many subject matter domains.

Finally, we now have a project underway in the small business program (in this case funded by the Office of the Secretary of Defense) that is building an authoring tool for case-based instruction, a form of instruction commonly used for management education and for some forms of advanced military education or training. This was inspired by some striking findings emerging from Gentner's basic research project – identifying more effective ways to do case-based instruction (Thompson, Gentner, & Loewenstein, 2000). This project, which also tries to incorporate aspects of the way cases are used at the Harvard Business School, is being conducted by Susann Luperfoy of Stottler-Henke Associates. Another such authoring tool project is focused exclusively on the use of cases in medical instruction. The courseware produced by these tools, however, will not be artificially intelligent tutoring systems.

Cognitive Task Analysis

It is generally agreed that doing cognitive task analyses needed for intelligent tutoring systems, for other training systems for jobs and tasks with significant cognitive components, and for the design of human system interaction is a major bottleneck area, still more art than engineering, requiring highly trained cognitive psychologists or cognitive scientists. Therefore we have been interested in supporting good ideas that show promise of moving this toward a more routine and reliable activity. But good ideas seem to be scarce. The most notable effort to date resulted in the CAT and CAT-HCI tools developed by Kent Williams (Williams, 2000), now at the University of Central Florida. These aid, support and record hierarchical cognitive task analyses, and were designed to go all the way down to the fine-grained production system level of analysis such as seen in Anderson's cognitive models and tutors. CAT-HCI incorporates information from the psychological research literature about execution times for many cognitive and overt actions and can be used to evaluate human computer interaction designs by predicting task execution designs. It can also predict expected learning/training times, relying on past ONR-sponsored research by Kieras (Kieras & Bovair, 1986). However, there is no magic in these tools. The analyst must still have insight into the cognitive operations used.

Another notable product in this area resulted from a NATO Study Group on Cognitive Task Analysis chaired by Jan Maarten Schraagen of TNO in the Netherlands. The original goal of the study group was to match cognitive task analysis problems with task analysis methods. This proved unfeasible because most so-called methods address only a small aspect of any given cognitive task analysis problem. Any successful cognitive task analysis requires orchestrating quite a number of these small methods. However, ONR sponsored an international workshop to address the issue. This resulted in a published book reviewing the state of the art, *Cognitive Task Analysis* (Schraagen, Chipman, & Shalin, 2000) . This effort further inspired an ONR supported small business contract for a digital library resource on cognitive task analysis. Primarily, this was intended to assist people in performing cognitive task analyses by giving them access to past cases of successful task analyses, so that they could see how problems similar to their own had been addressed. The contract was awarded to Aptima, Inc. As of this writing the resource can be accessed at ctaresource.com. It remains to be seen whether the community will support the continuing existence of this resource.

Major Applied Demonstration Projects

Still more applied than the work discussed above, are demonstration projects (6.3) that demonstrate the feasibility of advanced training technology applied to a specific training problem with a defined military customer. These are intended not to produce a completely finished training product but to reduce risk before an actual training system procurement. Often, however, training demonstrations are closer to the final product than is the case for other areas of R&D because it is desirable to conduct experimental evaluations of training effectiveness. An experimental training system must be quite finished and reliable before such experiments can be conducted. Typically, these projects are quite expensive, much more so than other projects mentioned above.

The IMAT/DSOT (Interactive Multi-sensor Analysis Trainer/Deployable Sonar Operator Trainer) line of work discussed in the paper by Wulfeck represents one of the largest, longest-running and successful investments of this type. This work received the very first Bisson award recognizing successful transitions of R&D efforts into practical fleet use, a competition of across all kinds of Navy R&D.

Another major project which received a high level investment for 3 years was the Advanced Embedded Training (AET) project (Zachary, Cannon-Bowers, Burns, Bilazarian, & Krecker, 1998). This was the culmination of the line of research on Tactical Decision Making Under Stress (TADMUS) carried out at the Navy training lab in Orlando, Florida. This system automated the team training techniques developed for use with the AEGIS anti-air warfare team. It is the first example I know about of an intelligent tutoring system

for a team, as opposed to an individual. It emphasized performance assessment by analyzing keystroke and trackball movements, eye tracking on the displays and speech recognition to determine whether the right information was being asked for or communicated at the right time. Detailed cognitive models of the team members, models actually capable of performing the tasks, were built to serve as a basis of comparison with trainee performance. These were intelligent, flexible “expert systems” that could handle many different scenarios, and the total project also included scenario authoring capability. Although there was a minimal capability for feedback during performance, the emphasis was on after-action review provided by a human instructor using information provided by the automated systems. AET was intended as a demonstration of what might be done in future ships such as DD-21 or DD-X. It remains to be seen whether this type of sophisticated training capability will be implemented in future ships. The cognitive models built in this research have also been used in some further research on training with synthetic team mates. The team training research behind the AET has also received a Bisson award for fleet transition.

The DC-TRAIN system mentioned above was another such project, with a much smaller budget. It is designed for training the damage control assistant, an officer who coordinates damage control efforts. DC-TRAIN is a complex system incorporating numerical simulations of fire, the spread of fire, flooding, etc., that are physically correct. In addition, it has an expert system (of admittedly somewhat limited scope) that knows how to do damage control. As in the AET system, the expert system’s recommended actions are used as a basis for evaluating trainee performance. Instructional intervention is limited to an after-action critique of the trainee’s performance. Like AET, DC-TRAIN has the capability for easy authoring of a large number of different damage scenarios. As the software system became larger and more complex (over 2 million lines of code), it developed reliability problems. However, these seem to have been overcome; nearly all were traced to minor errors in the database describing the ship that caused the numerical simulations to crash. A lesson learned was that a more qualitative and less sensitive simulation approach would probably be a better choice for a practical training system procurement. DC-TRAIN’s written output to the student also is not as graceful as it could be. Nevertheless, DC-TRAIN has been well received by a working group on damage control training, which judged it promising for addressing the “command and control” dimension of damage control training. The natural language debriefing dialog demonstrated under the Stanford Tutorial Dialog project discussed above makes it significantly more attractive. Despite some difficulties along the way, DC-TRAIN should probably be regarded as a successful, risk reducing concept demonstration (Bulitko, V. V. and Wilkins, D. C., 1999; Sniezek, J. A., Wilkins, D. C., and Wadlington, P., (2002); see also <http://www-kbs.ai.uiuc.edu> for many more reports.)

Virtual reality has also begun to pay off in practical demonstration projects. The first successful example was VE-SUB, a virtual reality simulator providing the opportunity to practice bringing submarines into harbor. This project has moved on from a demonstration project to the actual procurement of the training systems, with simulations of quite a number of different harbors. It is being implemented in all six U.S. submarine training centers. Advancing beyond this project was COVE (Conning Officer’s Virtual Environment) which integrated the virtual environment simulator with an artificially intelligent coach of ship driving. This was evaluated as effective and is moving on to practical implementation. Additional vehicle trainers of this general type are also being developed for the Marine Corps.

ONR’s training R&D is doing very well in providing effective training technology to the fleet, as the Bisson awards illustrate – only a few such awards have yet been made, and training R&D has received two of them. Many other practical training applications, such as the current large investments in computer based training and web-based training, also owe their feasibility to ONR research investments, even though that connection is often forgotten and rarely recognized.

References

- Albacete, P.L., & VanLehn, K. (2000) Evaluation of the effectiveness of a cognitive tutor for fundamental physics concepts. In L.R. Gleitman & A.K. Joshi (Eds), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society* (pp. 25-30). Mahwah, NJ: Lawrence Erlbaum Associates.
- Anderson, J. R. (1993) *Rules of the Mind*. Hillsdale NJ: Lawrence Erlbaum Associates. (see chapter 11).
- Anderson, J. R. & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum
- Anderson, J. R., Boyle, D. F. and Reiser, B. J. Intelligent tutoring systems. *Science*, 1985, 288, 456-462.
- Atkinson, R.K. (2003) An experimental evaluation of tutorials in problem solving (TiPS): A remedial mathematics tutor. Final report on ONR N00014-02-1-0191.
- Atkinson, R., Derry, S. J., Renkl, A. & Wortham, D. (2001). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70, 181-215
- Atkinson, R. K., Wortham, D. W., Derry, S. J., Jiang, N., & Gance, S. K. (1998). *Beyond static worked examples: Testing the efficacy of computer delivered dynamic examples*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.
- Byrne, M. & Anderson, J. R. (1997). Enhancing ACT-R's perceptual-motor abilities. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*, p. 880. Mahwah, NJ: Erlbaum.
- Bloom, B.S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13, 4-16.
- Bulitko, V. V. and Wilkins, D. C., (1999) "Automated Instructor Assistant for Ship Damage Control," *Proceedings of the Eleventh Conference on Innovative Applications of Artificial Intelligence*, IAAI-99, Orlando, July 18-22, 1999, 778-785.
- Chi, M.T.H., Bassok, M., Lewis, M.W., Reimann, P., & Glaser, R. (1989) Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 15, 145-182.
- Cho, B.I., Michael, J., Rovick, A., & Evens, M.W. (2000) An analysis of multiple tutoring protocols. *Proceedings Intelligent Tutoring Systems: ITS 2000*. Montreal. P. 212-221.
- Chong, R. S., & Laird, J. E. (1997). Identifying dual-task executive process knowledge using EPIC-Soar. In *Proceedings of the 19th Annual Conference of the Cognitive Science Society*. 107-112. Mahwah, NJ: Lawrence Erlbaum.
- Clancey, W. J. (1986) From GUIDON to NEOMYCIN and HERACLES in twenty short lessons: ONR final report 1979-1985. *AI Magazine*, 1986 (August), 40-187.
- Core, M.G., Moore, J.D., Zinn, C. (2003) The Role of Initiative in Tutorial Dialogue, Eleventh Conference of the European Chapter of the Association for Computational Linguistics (EACL), Budapest, Hungary, 2003.

Derry, S. J., Wortham, D., Webb, D., & Jiang, N. (1996). *Tutorials in problem solving (TiPS): Toward a schema-based approach to instructional tool design*. Paper presented at the annual meeting of the American Educational Research Association, New York, NY.

Di Eugenio, B. (2001) Natural-language processing for computer-supported instruction. *ACM Intelligence*, 12, p.22-33.

Doane, S.M., & Sohn, Y.W. (2000) ADAPT: A predictive cognitive model of user visual attention and action planning. *User Modeling and User Adapted Interaction*, 10, 1-45.

Evens, M.W. (2000) *Language Understanding and Generation in Complex Tutorial Dialogues*. Final report on grant N00014-94-1-0338. Chicago: Computer Science Department, Illinois Institute of Technology.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence*, 41, 1-63.

Forbus, K. D., Gentner, D., & Law, K. (1995). MAC/FAC: A model of similarity-based retrieval. *Cognitive Science*, 19(2), 141-205. (Abridged version to be reprinted in T. Polk & C. M. Seifert, (Eds.), *Cognitive Modeling*. Boston: MIT Press)

Forbus, K. & Whalley, P. (1994) Using qualitative physics to build articulate software for thermodynamics education. *Proceedings of AAAI-94*, Seattle.

Fox, B. A. (1993). *The Human Tutorial Dialogue Project: Issues in the design of instructional_systems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Fry, J., Ginzton, M., Peters, S., Clark, B. & Pon-Barry, H. (2001) Automated tutoring dialogues for training in shipboard damage control. In Proceedings 2nd SigDial Workshop on Discourse and Dialogue.

Gentner, D., Bowdle, B., Wolff, P., & Boronat, C. (2001). Metaphor is like analogy. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), (2001). *The analogical mind: Perspectives from cognitive science* (pp. 199-253). Cambridge, MA: MIT Press.

Gott, S. P. (1989) Apprenticeship instruction for real-world tasks: The coordination of procedures, mental models, and strategies. In: Rothkopf, E. Z. (Ed.), *Review of research in education*, 15, (pp. 97-160). Washington, DC: American Educational Research Association.

Graesser, A.C., Person, N.K., & Magliano, J.P. (1995) Collaborative dialogue patterns in naturalistic one-to-one tutoring. *Applied Cognitive Psychology*, 9(6), 495-522.

Graesser, A.C., VanLehn, K., Rose, C., Jordan, P. & Harter, D. (2001) Intelligent tutoring systems with conversational dialogue. *AI Magazine*, 22(4), 39-51.

Hume, G., Michael, J., Rovick, A., & Evens, M. (1996) Hinting as a tactic in one-on-one tutoring. *The Journal of the Learning Sciences*, 5(1), 23-47.

Just, M. A., Carpenter, P. A., & Varma, S. (1999). Computational modeling of high-level cognition and brain function. *Human Brain Mapping*, 8, 128-136.

Just, M. A., & Varma, S. (2002). A hybrid architecture for working memory: Reply to MacDonald and Christiansen (2002). *Psychological Review*, 109, 55-65.

Kieras, D.E., & Bovair, S. (1986) The acquisition of procedures from text : A production system analysis of transfer of training. *Journal of Memory and Language*, 25, 507-524.

Kieras, D.E. & Meyer, D. E. (1997) An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human Computer Interaction*, 12, 391-438.

Koedinger, K.R., Anderson, J.R., Hadley, W.H., & Mark, M. (1995) Intelligent tutoring goes to school in the big city. In: J. Greer (Ed) *Artificial Intelligence in Education*, 1995, p.421-428 Proceedings of AI-ED 95, Washington, DC: August 16-19.

Landauer, Thomas K.; Foltz, Peter W.; Laham, Darrell (1998) An introduction to latent semantic analysis. *Discourse Processes*, 25(2-3) 259-284

Lesgold, A., Chipman, S., Brown, J.S. & Soloway, E. (1990) Intelligent training systems, *Annual Review of Computer Science*, 4, 383-94.

Marshall, S. P. (1995). *Schemas in Problem Solving*. Cambridge, England: Cambridge Press.

McCarthy, J.E., Pacheco, S., Banta, H.G., Wayne, J.L., & Coleman, D.S. (1994, November) The Radar System Controller Intelligent Training Aid. Paper presented at the 16th Interservice/Industry Training Systems and Education Conference, Orlando, FL.

Meyer, D.E. & Kieras, D. E. (1997a) A computational theory of executive cognitive processes and multiple-task performance: Part 1 Basic Mechanisms. *Psychological Review*, 104, 3-65.

Meyer, D.E. & Kieras, D. E. (1997b) A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory period phenomena. *Psychological Review*, 104, 749-791.

Meyer, D. E. & Kieras, D. E. (1999) Precis to a practical unified theory of cognition and action: Some lessons from computational modeling of human multiple-task performance. In D. Gopher & A. Koriat (Eds) *Attention and performance XVII* (pp. 15-88). Cambridge, MA: MIT Press.

Moore, J.D., Lemaire, B., & Rosenblum (1996) Discourse generation for instructional applications: Identifying and exploiting relevant prior explanations. *Journal of the Learning Sciences* 5(1), 49-94.

Newell, A. (1990) *Unified Theories of Cognition*. Cambridge, MA: Harvard University Press.

Schofield, J. W. (1995) *Computers and classroom culture*. Cambridge University Press.

Schraagen, J.M., Chipman, S.F. & Shalin, V. (Eds) (2000) *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Shelby, R.N., Schulze, K.G., Treacy, D.J., Wintersgill, M.C., & VanLehn, K. (2001). The Andes Intelligent Tutor: An Evaluation. Paper presented at the Physics Education Research Conference, Rochester, NY, July 21-25.

Sohn, M.-H., Douglass, S. A., Chen, M.-C., & Anderson, J. R. (submitted). Unit task execution with a dynamic display in a complex problem-solving situation.

Snizek, J. A., Wilkins, D. C., and Wadlington, P., (2002) Training for Crisis Decision Making: Psychological Issues and Computer-Based Solutions, *Journal of Management Information Science*. Volume 18, Number 4, 147-168.

Thompson, L., Gentner, D., & Loewenstein, J. (2000). Avoiding missed opportunities in managerial life: Analogical training more powerful than individual case training. *Organization Behavior and Human Decision Processes*, 82(1), 60-75.

Towne, D. M. (1987) The Generalized Maintenance Trainer: Evolution and Revolution. In: William B. Rouse (Ed) *Advances in Man-Machine Systems Research, Volume 3*. Greenwich, CN: JAI Press, p. 1-63.

Tuttle, K., & Wu, C. (2001) *Thermodynamics: A computer based approach*. Dubuque, Iowa: Kendall/Hunt. ISBN 0-7872-8902-7.

Vasandani, V. (1991) Intelligent Tutoring for Diagnostic Problem Solving in Complex Dynamic Systems. Atlanta, Georgia: Georgia Institute of Technology, School of Industrial and Systems Engineering, Center for Human-Machine Systems Research, TR CHMSR-91-4.

Wenger, E. (1987) *Artificial Intelligence and Tutoring Systems: Computational and Cognitive approaches to the communication of knowledge*. Los Altos, CA: Morgan Kaufman.

Wetzel, C.D., Radtke, P.H., & Stern, H.W. (1994) *Instructional Effectiveness of Video Media*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Williams, K.E. (2000) An Automated Aid for Modeling Human-Computer Interaction. In J.M. Schraagen, S.F. Chipman, & V. Shalin (Eds) (2000) *Cognitive Task Analysis*. Mahwah, NJ: Lawrence Erlbaum Associates.

Wortham, D. W. (1996). *Testing a learning environment based on a semantic analysis of math problems*. Master's Thesis, Department of Educational Psychology, University of Wisconsin-Madison.

Wortham, D. W., Webb, D. C., & Atkinson, R. K. (1997). *Effects of two curricula on proportional reasoning, problem solving, and strategy use*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Wu, C. (2002) *Intelligent Computer-Based Elementary Applied Thermodynamics*, Dubuque, Iowa: Kendall/Hunt, ISBN 0-7872-9388-1.

Wu, C. (2003) *Intelligent Computer Based Engineering Thermodynamics*. New York: Nova Science. ISBN 1-59033-359-4.

Zachary, W., Cannon-Bowers, J., Burns, J., Bilazarian, P., & Krecker, D. (1998) An advanced embedded training system (AETS) for tactical team training. In B.P. Goettl, H.M. Halff, C.L. Redfield, & V.J. Shute (Eds) *Intelligent Tutoring Systems: Proceedings of the 4th International Conference, ITS '98*, 544-553. Berlin: Springer Verlag.

Zachary, W., Ryder, J., Hicinbothom, J., Santarelli, T., Scolaro, J., Szczeplkowski, M., and Cannon-Bowers, J. (1998) Simulating Behavior of Tactical Operators and Teams Using COGNET/GINA. In: Proceedings of

the 7th Symposium on Computer-Generated Forces and Behavioral Representation. Orlando, Fl: Institute for Simulation and Training.

Zinn, C., Moore, J.D., and Core, M.G., (2002a) A 3-tier Planning Architecture for Managing Tutorial Dialogue, *Intelligent Tutoring Systems, 6th. Intl. Conference, ITS 2002*, Springer LNCS 2363, pages 574-584, Biarritz, France, June 2002.

Zinn, C., Moore, J.D., and Core, M.G., (2002b) A 3-tier Planning Architecture for Managing Tutorial Dialogue, System Demonstration, page 200, *Proceedings of the sixth workshop on the semantics and pragmatics of dialogue (EDILOG 2002)*, Edinburgh, UK.

Zinn, C., Moore, J.D., and Core, M.G, (in press, 2003) Intelligent information presentation for tutoring systems. In: O. Stock & M. Zancanaro (Eds): *Intelligent Information Presentation*, Dordrecht, NL: Kluwer Academic Publishers.

Appendix A

Transcript of a Tutorial Dialog for After-action Review in DC-TRAIN

Tutor: Hello, we are about to review your session from yesterday.

Tutor: All right, are you ready?

Student: Yes.

Tutor: In this session there were 5 primary crises and 1 secondary crisis.

Tutor: 5 of the 6 crises were resolved.

Tutor: Most of your actions were correct, however you made several mistakes.

Tutor: OK, let's discuss a few of those crises.

Tutor: The first crisis was a fire here in Crew Living space No. 1. [centers on this compartment]

Tutor: This fire was not contained early enough.

Tutor: As a result, the fire spread to the adjacent Combat system equipment room No. 2. [highlights 2nd compartment]

Tutor: OK, let's review how to handle this crisis. [removes highlight from 2nd compartment]

Tutor: What's the first thing you should do when the fire alarm sounds?

Student: Send repair two to investigate.

Tutor: Yes, that's right.

Tutor: However, you failed to do this during the session.

Tutor: This is a mistake you made more than once.

Tutor: Remember, alarms need to be verified first.

Tutor: OK, suppose that there is a fire.

Tutor: What's the next step?

Student: Set fire and smoke boundaries.

Tutor: Yes.

Tutor: That's what you did during the session.

Tutor: OK, what else needs to be done?

Student: The repair team should fight the fire.

Tutor: No.

Tutor: You missed a step.

Tutor: Remember, the crisis might interfere with other ship systems.

Tutor: Let's try again.

Tutor: What else needs to be done?

Student: I don't know.

Tutor: Well.

Tutor: You should order repair five to mechanically and electrically isolate the compartment.

Tutor: But you didn't do this during the session.

Tutor: This is a mistake you made more than once.

Tutor: Remember, mechanical and electrical isolation protects crucial ship systems.

Tutor: All right, suppose isolation has begun and boundaries have been set.

Tutor: What's the next step?

Student: Fight the fire.

Tutor: Yes, exactly.

Tutor: That's what you did.

Tutor: Suppose the fire has been overhauled.

Tutor: Let's now discuss the resulting smoke. [displays gray color in compartment]

Tutor: What should you do first?

Student: Send repair two to desmoke the compartment.

Tutor: Almost.

Tutor: You should send a different repair team.

Tutor: Which team should you send instead?

Student: Repair five.

Tutor: Yes, exactly.

Tutor: You should order repair five to desmoke the compartment.

Tutor: But you sent the wrong repair team during the session.

Tutor: Repair 2 was overtasked, so repair five should have been sent.

Tutor: OK, let's now discuss the flooding that resulted from your firefighting efforts.

Tutor: What's the first thing you should do in response to this crisis?

Student: Can we move on please?

Tutor: OK, let's talk about something else

Tutor: The next crisis was a flood here in Engine room No. 2. [displays compartment]

Tutor: Let's review how to handle this crisis.

Tutor: What's the first thing you should do when the flood alarm sounds?

Student: Send repair two to investigate.

Tutor: Almost.

Tutor: Let's review the jurisdiction of repair lockers.

Tutor: The flood alarm was here in Engine room No. 2.

Tutor: Please click on the location of the nearest repair locker in the left display above.

Tutor: Let me know when you're ready to continue.

[Student clicks on correct compartment]

Student: Please continue.

Tutor: Yes, that's right.

Tutor: Which repair locker is that?

Student: Repair five.

Tutor: Yes, that's right.

Tutor: You should order repair five to investigate the compartment.

Tutor: But you sent the wrong repair team during the session.

Tutor: All right, suppose that there is a flood.

Tutor: What should you do next?

Student: Set flood boundaries.

Tutor: You're right.

Tutor: However, you failed to do this during the session.

Tutor: This is a mistake you made more than once.

Tutor: Remember, setting flood boundaries helps keep floods from spreading.

Tutor: OK, what else needs to be done?

Student: I should isolate the compartment.

Tutor: Yes.

Tutor: However, you failed to do this during the session.

Tutor: We discussed this same mistake earlier.

Tutor: Remember, mechanical and electrical isolation protects crucial ship systems.

Tutor: OK, suppose isolation has begun.
Tutor: What's the next step?
Student: Can we move on to the next topic please?
Tutor: OK, let's talk about something else
Tutor: Would you like to discuss how you managed the firemain?
Student: Yes.
Tutor: What's the first thing you should do after General Quarters has been called?
Student: Order the DCCO to start a firepump.
Tutor: No, that's not right.
Tutor: You should have done something else first.
Tutor: You need to ask permission to do something.
Tutor: Now try again.
Tutor: What's the first thing you should do after General Quarters has been called?
Student: I don't know.
Tutor: Well.
Tutor: You should request permission from the EOOW to start a firepump.
Tutor: But you didn't do this during the session.
Tutor: Remember, you need the EOOW's permission before starting a firepump.
Student: Why?
Tutor: Well.
Tutor: This is specified in damage control doctrine.
Tutor: Would you like me to display the relevant section of doctrine?
Student: Yes.
Tutor: OK. [doctrine is displayed]
Tutor: Let me know when you're ready to continue.
Student: Please continue.
Tutor: OK, suppose you've received permission to start a firepump.
Tutor: What should you do next?
Student: Order the DCCO to start a firepump.
Tutor: Yes, that's right.
Tutor: You did this during the simulation.
Tutor: OK, let's review the most important points from this session.
Tutor: You performed well in the areas of investigating, containing, and eliminating crises.
Tutor: We discussed your errors relating to isolation, and permission to start a firepump.
Tutor: Remember, it's important to mechanically and electrically isolate a compartment in order to protect crucial ship systems.
Tutor: Also, you must request permission from the EOOW before starting a firepump.
Tutor: OK, that ends our review.
Tutor: Would you like to review this session again?
Student: Nope.
Tutor: Goodbye.

